**贝叶斯--实验说明**

**A．实验任务题目：**

1. 利用贝叶斯算法进行数据分类操作，并统计其预测正确率。

2. 随机产生10000组正样本和20000负样本高斯分布的数据集合（维数设为二维），要求正样本：均值为[1;3]，方差为[2 0;0 2]；负样本：均值为[10;20]，方差为[10 0;0 10]。先验概率按样本量设定为1/3和2/3.分别利用最小错误概率贝叶斯分类器和最小风险概率贝叶斯分类器对其分类。（假设风险程度正样本分错风险系数为0.6，负样本分错风险为0.4，该设定仅用于最小风险分析）

3. 编写一个贝叶斯分类器。输入为：均指向量、先验概率、协方差矩阵、输入学习数据X,测试数据类别XLABEL,测试数据Y.输出为Y对应的类别。（选做）。

**B．编程语言：**

Python、Matlab、C/C++、Java、R语言均可

**C．实验数据介绍：**

**汽车评价数据集**

共1728个数据，每个数据特征为6维，分为4类，类别标记为unacc，acc，good，V-good

四个类别标记分别表示汽车性价比等级（由低到高）

unacc：1210个

acc：384个

good：69个

V-good：65个

6个特征分别为：（6个属性）

1. buying （取值：v-high、high、med、low） 表示购买价格
2. maint （取值： v-high、high、med、low） 表示维修价格
3. door （取值：2、3、4、5-more） 车门数量
4. Persons （取值：2、4、more） 可容纳人数
5. Lug\_boot （取值：small、med、big） 行李箱大小

Safety （取值：low、med、high） 安全系数

链接：http://archive.ics.uci.edu/ml/datasets/Car+Evaluation

**D．实验完成要求：**

1. 仔细阅读并了解实验数据集；
2. 使用任何一种熟悉的计算机语言（比如 C，Java或者MATLAB）实现朴素贝叶斯算法；
3. 利用朴素贝叶斯算法在训练数据上学习分类器，训练数据的大小分别设置为：前100个数据，前200个数据，前500个数据，前700个数据。前1000个数据，前1350个数据；
4. 利用测试数据对学习的分类器进行性能评估；
5. 演示实验，提交代码，统计分析实验结果并上交实验报告；